RAND

The Role of Federal Resources in Closing the Achievement Gaps of Minority and Disadvantaged Students

David Grissmer and Ann Flanagan

DRU-2501-EDU

February 2001

RAND Education

THE ROLE OF FEDERAL RESOURCES IN CLOSING THE ACHIEVEMENT GAPS OF MINORITY AND DISADVANTAGED STUDENTS

David Grissmer and Ann Flanagan

RAND

INTRODUCTION

In their landmark volume, Jencks and Phillips (1998) suggested that reducing the gap in scores between black and white students would do more to move America toward racial equality than any other alternative. They argue that eliminating the gap might eventually solve many of the others issues that separate the races: the need for affirmative action policies in colleges and jobs, the gap in wages, and perhaps even segregated communities. When recommending policies that could help achieve gap reduction, they suggest that the strongest empirical evidence exists for class size reductions and raising the quality of teachers by obtaining teachers with higher test scores as two "expensive" options. They also believe that policies prior to kindergarten such as pre-kindergarten and Headstart programs that are oriented toward teaching cognitive skills are necessary.

They also emphasize the difficulty of finding politically viable solutions. Decisive political support for programs that only benefit blacks is unlikely, and so benefits for some whites seems necessary to build a viable political coalition. Jencks and Phillips (1998) also were pessimistic that expanding federal programs could play a significant role in solving the problem because of the distrust of federal initiatives in education.

In this paper we will suggest that a federal resource role is critical to addressing the problem, and that specific programs that are appropriately targeted could be effective and efficient. While significant political issues of federal involvement are still present, the recent election and the federal surplus has probably made a broader federal role in education more likely– although still difficult.

Finally, we will explore the possibility of a political coalition that could form to address the significant resource issues in U.S. education.

The paper is based on several premises that we first discuss. These include the following:

- Compelling empirical evidence now exists that appropriately targeted expenditures to specific programs can raise achievement
- Although the evidence is less compelling, there is significant evidence that additional money spent on disadvantaged and minority students is more effective and efficient than on their more advantaged counterparts.
- Most of the inequality in educational spending is between states, not within states. Even if court suits were successful across states in reducing inequity within each state, most of the inequality in spending would still be present due to between state differences
- Although low scoring students are disproportionately minority and disadvantaged, about an equal number are white, black and Hispanic, and are disproportionately located in the south and central cities, suggesting a possible, but certainly hard to form, political coalition.
- Federal resources would be required to address between state inequality in resources, perhaps through a restructured and greatly expanded Title I type program that is better targeted By locality and focuses expenditures on programs with strong empirical support
- Federal resources can also be used to increase teacher quality nation-wide, particularly in localities and subjects in short supply, through appropriately structured scholarships, loan forgiveness and "GI Bill" type incentives.

We first discuss the evidence for the first three premises. Then we present an analysis using a composite sample of nine national NAEP tests for public school students that describes the average achievement scores and population percentages by race, region and locality (central city, suburban and rural). This analysis also identifies the characteristics

of the lowest scoring quintile of students by race, region and locality. Finally, we suggest programs where targeted federal resources could impact the achievement of lower scoring students through improving teacher quality and providing more resources for lower class size and pre-kindergarten.

RESOURCES AND INEQUALITY

As long as educational research was unable to provide consistent and replicable measurements on the effects of more and different uses of resources on educational outcomes, policymakers and school systems lacked perhaps the most important ingredient to improve the quality of education. Also missing were measurable educational objectives and internal incentives focused on meeting those objectives. In such an environment it is easy to understand the pessimism of the late 1980s and early 1990s about whether public education could be reformed. Effective and predictable utilization of resources that allow measurable results is a necessary condition for reforming education. But good research alone is not sufficient. It must be accompanied by internal incentives and measurable goals that motivate its use and ongoing evaluation.

Prior to the mid 1990s, apparent weak and inconsistent empirical evidence on the effects of resources fostered one view that public school systems could not turn additional resources into better educational outcomes (Hanushek, 1989; Hanushek, 199x, Hanushek, 199x, Hanushek, 199x). These studies utilized a "vote counting" approach applied to published measurements. However, two re-reviews utilizing the same set of studies came to more positive conclusions. Hedges et al. (1994) used meta-analysis that tested specific statistical hypothesis about the distribution of previous measurements, and concluded that resources showed a net positive relationship to outcomes. However, wide variation in measurements still made it difficult to apply to policy. Krueger (2000) also re-reviewed Hanushek's studies and concluded that the choice of including multiple measurements from individual studies, and the possible selectivity involved in publishing multiple measurements was unduly influential in making the results look weak and inconsistent.

Greenwald et al. (1996) utilized tighter quality criteria, but utilized studies from a wider set of journals, and used meta-analysis to reach a conclusion that most resources showed net positive relationships with outcomes. However, all reviews show an uncomfortably wide variance in measurement results.

Ferguson and Ladd (1996) suggests that this inconsistency in previous measurements may be due to different and flawed specifications and assumptions across models. Grissmer et al. (2000) suggests, in addition to model specification flaws, that differential bias across levels of aggregation and differential effects across types of students may partially account for such differences. In their literature review and analysis of the state NAEP data, they conclude that the weight of evidence is shifting toward an hypothesis that resources have their largest and most efficient impact on disadvantaged and minority students.

They cite an emerging consistency of evidence from experimental data, explanations for previous non-experimental analysis and the historical pattern of score gains from 1970-1996. In the latter case, the larger gains in achievement by minority and disadvantaged white students with smaller or no gains by advantaged white students seems difficult to explain outside of the additional resources targeted to these students or for programs that differentially benefit these students such as pupil-teacher reductions. Ferguson (1998) Grissmer et al. (1998) and Krueger (1999) all suggest that pupil-teacher reductions are consistent with explaining part of the large black score gains.

The strongest evidence for effects of particular resources and possible differential effects for disadvantaged and minority students comes from experimental or quasi-experimental research on class size (Finn and Achilles, 1999; Krueger, 2000; Molnar et al., 1999; Nye et al., 1999; and Nye et al., forthcoming). Summarizing this research, Grissmer et al, 1999 concludes that the studies show effects that are large and sustained into later grades for students with small class size in K-3. The measured effects also are larger for minority and disadvantaged students. Hanushek (1999) details the possible flaws in the experiments that could affect these results. However, empirical

tests exploring the sensitivity of results to flaws in the experimental execution have not shown sample attrition, leakage between test and control groups, teacher assignments and non-participation to significantly affect results(Krueger, 2000; Nye et al., 2000; Nye et al., forthcoming).

Perhaps as importantly, the Tennessee experimental results seem to challenge commonly used model specifications, and might help explain the inconsistency of previous non-experimental results. The pattern of results suggest that inclusion of resources since school entry is important in models, and that use of previous year's test scores to control for both missing family characteristics and earlier resources is problematic. Since these are common flaws in specifications in previous studies, these results may help explain the inconsistency of previous results.

If we accept the premise that money can make the most difference for disadvantaged and minority students, the next question involves the apparent inequity in funding for advantaged and disadvantaged students. At one level, this inequity has been repeatedly challenged within states through the courts (Ladd et al., 1999). Starting with the constitutional challenge in California in 1971 to the inequity across school districts in per pupil expenditures, finance reform litigation has occurred in 43 states (Evans et al., 1999). Courts have overturned systems in 19 states, and upheld systems in 20 states with 4 still pending.

Some states have made significant changes in the way education is financed by decreasing or eliminating reliance on the property tax in order to reduce inequality in per pupil expenditures. The results of such judicial intervention overturning existing financing systems seems to be increases in total state funding for education, significant narrowing of variance in spending across districts with lower spending districts being the main beneficiaries (Evans et al., 1999). Their analysis also showed black and white students receiving roughly equal net spending increases due to judicial intervention.

Only a few states have substantially eliminated variation in per pupil expenditures across districts, and judicial intervention seems to be required to do so. While it seems politically impossible to reduce

expenditures for more advantaged students in order to spend more on disadvantaged students, a more achievable goal has been set out as insuring an "adequate" education for all students. So courts are increasingly facing adequacy rather than equity suits (Ladd and Hansen, 1999). Such suits will continue across the states, and perhaps be more easily won if research consistently points to larger effects of resources on disadvantaged students. However, the problem being addressed by these actions- within state inequality in spending—is not the primary source of inequity in educational spending. Between state inequality in spending accounts for almost two-thirds of the variance in spending while within state inequality accounts for only a third (Evans et al, 1999). Equalizing spending in all states would still leave significant variation in expenditures across the nation. For instance in 1995-96, adjusting for cost of living differences, Mississippi spent $4,900 per pupil, while New Jersey spent $9,090 per pupil. Low spending states are disproportionately in the south, while the northeast and midwest have the highest spending states. Current federal spending narrows this gap only slightly. For instance, Mississippi in 1995-96 received approximately $590 per pupil, while New Jersey received $340 per pupil. After cost of living adjustments, federal spending narrows the gap by about $400 dollars a pupil, leaving a gap of $4,000 per pupil.

Setting aside the issue of the politics of more federal spending directed toward insuring adequacy across states, the question arises as to where resources should be directed. If lower scoring students represent the most effective use of resources to increase national achievement, the question is where are lower scoring students located. The next section analyzes the national NAEP scores by race, region and locality to address this question.

EXPLORING WHITE AND BLACK NAEP SCORES BY REGION, LOCALITY AND STATE

The Data

Some of the distinct patterns present in U.S. achievement scores can only be seen by disaggregating by race, region and locality. The

racial differences are the most pronounced, and since black students are not distributed proportionately by region or locality, the aggregate region and locality scores can largely reflect the racial distributions. There are also significant differences by region in the SES status of students in central cities, suburbs and rural areas.

The problem in looking at these patterns is that the sample size of tests administered nationally are usually insufficient to produce robust samples at such disaggregated levels. We have combined nine national NAEP tests(main assessments) given from 1990-1998 (See Table 1). These tests were given to representative national samples including public and private school students. Our analysis only includes public school students at 4th and 8th grade. The public school sample size per test varied from around 3000 to 9000, with the total combined sample being approximately 52,000. The nine tests include three 4th grade reading tests, three 8th grade math tests, two 4th grade math tests and one 8th grade reading test. So the combined sample includes a composite picture of reading and math achievement at 4th and 8th grade during the 1990s.

Year	Grade	Subject	Public School Sample
1990	8th	Math	2879
1992	8th	Math	6033
1992	4th	Read	5045
1992	4th	Math	5641
1994	4th	Read	6030
1996	4th	Math	5215
1996	8th	Math	5590
1998	4th	Read	6300
1998	8th	Read	9091

Table 1. Nine NAEP Main Assessments Given From 1990-1998

The NAEP main assessments are not considered "basic skills" tests since in addition to multiple choice items, they include constructed response items requiring responses from a few sentences to a few paragraphs. NAEP data collection takes students approximately 90 minutes to complete for a given subject. Matrix sampling of questions is used to allow testing a broad range of knowledge while limiting the time each student is tested. Bib spiraling of questions insures that effects from the placement of questions within booklets and grouping

sets of questions are minimized. Two types of exclusions from testing are allowed: Limited English Proficiency (LEP) and Individualized Education Plan/Disabled (IEP/DS). Approximately 1-2 percent of students nationally are excluded for LEP and about 4-6 percent for IEP/DS.

The NAEP samples are stratified in order to allow the reporting of scores for three mutually exclusive location types: central city, urban fringe/large town and rural/small town. Central cities are defined as the central cities of the Standard Metropolitan Statistical areas (SMSA's) as defined by the Office of Management and Budget (OMB). Suburban areas include densely settled places and areas within SMSA's classified as urban by OMB, but outside central cities and large towns (non-SMSA's) with a population more than 25,000. Rural areas include small towns outside SMSA's with a population of less than 25,000 and all places with population smaller than 2,500 outside SMSA's. In our sample approximately 30 percent of K-8 students were in central cities, 50 percent in suburban areas and 20 percent in rural areas.

Results

Figure 1 shows a black-white gap of almost a full standard deviation for the composite sample. Regional and locality average scores

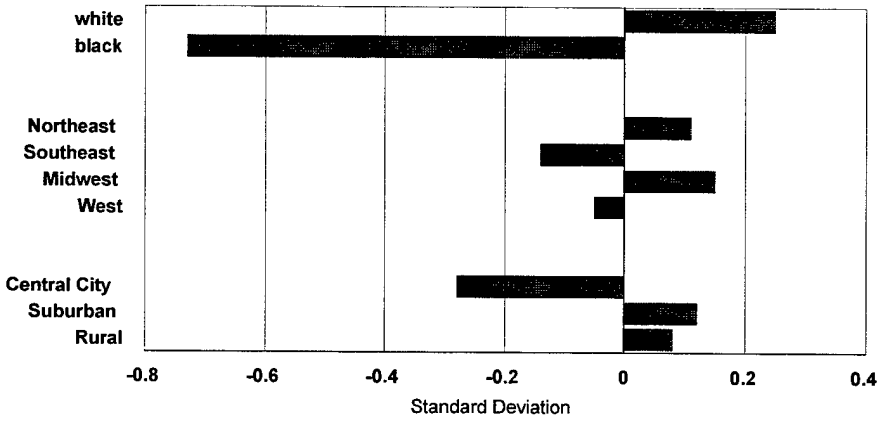


Figure 1. Average Scores Across Nine NAEP Tests for Public School Students by Race, Region and Locality

show higher scores in the northeast and midwest and in suburban and rural areas. Table 2 shows that the black-white gap is somewhat higher

for math than reading. Before presenting the more detailed data, we will summarize the results to make it easier to follow.

Type of test	Average White Score	Average Black Score	Gap in Standard Deviation
4th math	228.2	195.2	1.08
4th read	223.5	190.4	0.88
8th math	276.4	238.2	1.11
8th read	269.4	241.6	0.86

Table 2. Black-White Gap by Type of Test

Summary of locational data

The data show that there is a three tiered pattern of achievement in the U.S that can be identified by region and locality: 1) high scores in rural and suburban areas in the northeast and midwest; 2) average scores in rural and suburban areas of the south and west; and 3) low scores in central cities in all regions.

The most striking evidence comes from the midwest and northeast. About one-third of U.S. students live in the rural and suburban areas of the northeast and midwest, and these students have scores that place them near the top of international scores. White students in these areas - almost 30 percent of students nationally - would very likely place first in international competition.

However, central city students in these two regions are among the lowest scoring in the nation. Black and white students in central cities in these two regions have among the lowest scores when compared to black and white students in other regions or localities. The gap between central city and suburban scores are largest in these two regions. The states with the largest black-white score gap are Minnesota, Connecticut, Wisconsin and Michigan - all relatively high spending states in the northeast or midwest.

The rural and suburban areas of the south and west form a second tier and have average to below average scores. The third tier is central city scores in all regions with the lowest scores in the nation.

Black students perform best in the rural and suburban areas of the northeast, midwest and south. Black students perform worst in the west and all central cities regardless of region. White students do best in the rural and suburban areas of the northeast and midwest, and worst in central cities in all regions and the rural south. Interestingly, black scores in the rural south are among the highest black scores.

If we accept the premise that lowest scoring students provide the most effective investment, the location and characteristics of the lowest scoring students in the nation becomes important. We use the 20th percentile and below to define low scoring students. We find that there are approximately equal numbers of white, black and Hispanic students in the lowest quintile. The black and Hispanic students in the lowest quintile are highly concentrated. About 80 percent of blacks in the lowest quintile are in the south or central cities in the midwest and northeast. About 70 percent of Hispanics in the lowest quintile are in central cities or suburbs of the south and west. White students in the lowest quintile are more widely distributed. The largest groups are in the suburban areas of all regions and the rural south.

The south contains the largest proportion of low scoring students- about 42 percent of low scoring students are in the south. The largest concentration of low scoring students live in and around southern cities. One of three low scoring students lives there. The south and west together have 2 of 3 low scoring students. The central cities of the midwest and north have only 17 percent of the nation's low scoring students.

Examining Disaggregated Scores

Table 3 shows an interesting pattern when locality is disaggregated by region. For instance the data show that students in the suburban northeast constitute about 11.6 percent of students nationally and have the highest score of any group- .32 standard deviation above the national average. The rural and suburban areas of the northeast and midwest are grouped together at the top of the score distribution, significantly above any other group. Students in these high scoring

areas represent about one-third of students nationally (from column three).

To get an idea of how well students in these areas perform, we have made an approximate translation of their scores onto an international distribution. Phelps et al, 1996 develops a comparison of international scores to state NAEP scores for 8th grade math. In this comparison, the average scores for the state of Iowa place it slightly below Taiwan, the highest scoring country. The average score of Iowa students on the nine

Locality	Region	Percent of Student Population	Average Test Score (Standard Deviation)
Suburban	Northeast	11.6	0.32
Rural	Northeast	3.6	0.29
Suburban	Midwest	12.4	0.28
Rural	Midwest	5.4	0.26
Rural	West	3.3	0.07
Suburban	West	11.6	-0.03
Suburban	Southeast	14.8	-0.05
Rural	Southeast	7.3	-0.14
Central city	West	7.0	-0.15
Central city	Midwest	6.6	-0.22
Central city	Southeast	10.8	-0.30
Central city	Northeast	5.5	-0.46

Table 3 Average scores and percent of student population by region and locality

tests that are also administered across state samples (the state NAEP tests) was .30 standard deviation. Thus the average scores in rural and suburban areas of the Northeast and Midwest- about one-third of U.S. students-- would place their students very near the top of international score distributions.

Students in rural and suburban areas of the west and south fall in the middle of the distribution, while the scores of central city students from all regions are the lowest scores. Interestingly, the central city scores in the northeast are the lowest while suburban scores in the northeast are highest. The gap between central city and suburban scores in the northeast is .78 standard deviation-- by far the largest central city/suburban gap across regions.

Table 4 shows the complete disaggregation by race, region and locality. White students in rural and suburban areas of the northeast and midwest-- the highest scoring groups-- would likely rank at the top of international distributions. These students are approximately 30 percent of students nationally. White students scoring lowest live in central cities outside the south or in the rural south and west.

Race	Locality	Region	Student Population Percentage	Average Test Score (Standard Deviation)
White	Suburban	Northeast	9.4	0.47
White	Suburban	Midwest	11.0	0.37
White	Rural	Northeast	3.3	0.37
White	Rural	Midwest	5.1	0.31
White	Suburban	Southeast	10.1	0.23
White	Suburban	West	8.5	0.18
White	Central city	Southeast	5.1	0.17
White	Rural	West	2.9	0.16
White	Central city	West	4.6	0.15
White	Central city	Midwest	4.1	0.14
White	Central city	Northeast	2.2	0.03
White	Rural	Southeast	5.6	0.00
Black	Suburban	Northeast	1.4	-0.38
Black	Suburban	Midwest	0.8	-0.49
Black	Rural	Southeast	1.2	-0.65
Black	Rural	Northeast	0.2	-0.68
Black	Rural	Midwest	0.1	-0.71
Black	Suburban	Southeast	3.1	-0.76
Black	Central city	Southeast	3.7	-0.79
Black	Central city	Midwest	1.9	-0.79
Black	Central city	West	0.6	-0.81
Black	Central city	Northeast	2.2	-0.84
Black	Suburban	West	0.7	-0.93
Black	Rural	West	0.1	-0.99

Table 4. Average NAEP scores and population percentages by race, region and locality.

The average scores of white students in all areas are significantly higher than for any black student group. However, the pattern of scores is similar in many ways across races. The two highest scoring groups for both black and white students live in suburban areas in the northeast and midwest. The rural areas of the northeast and midwest are also near

the top of both black and white scores. Central city students in the northeast of both races are near the bottom. Western students in any locality for both black and white students tend to rank lower, but particularly for blacks. Of all regions, black students score lowest in the west. An interesting contrast is that rural southern scores are among the highest black scores, but the lowest white scores.

White student scores show a different locality pattern across regions. In the northeast and midwest, there are large gaps between central city and both suburban and rural scores. In the west, scores in each locality are virtually the same, while in the south, the central city and suburban scores are close, but both substantially higher than white rural scores– the area of lowest white scores.

White and Black Scores by State

Ideally, disaggregation of scores by race and locality could be done by state rather than region. Regional patterns can hide important differences across states within a region. States are the main driver of educational policy, and the pattern of black and white scores by locality within states could support a more detailed analysis including family characteristics and resources. Unfortunately the sample sizes of state NAEP scores are limited to about 2500, and combining samples across tests is more difficult since many states either did not participate or participated in a limited number of tests.

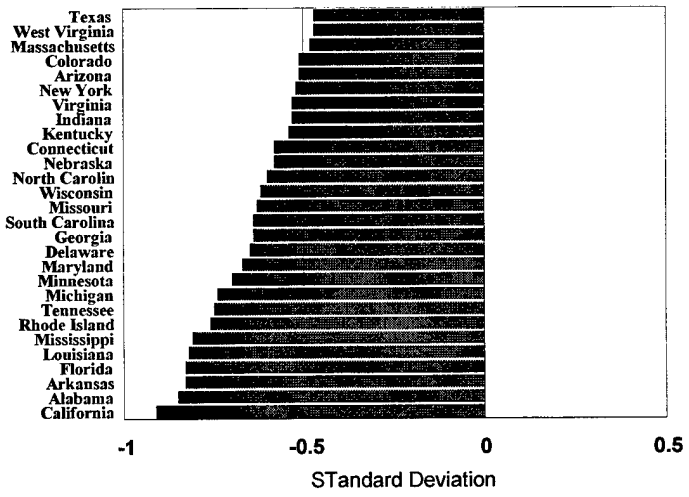


Figure 2 Average Black Scores by State Across Four State NAEP Tests(1992-1996)

We have combined the samples of four state NAEP tests all taken by 28 states. The four tests include the 1992 4th grade reading and math, the 1996 4th and 8th grade math. Figure 2 shows the average black scores by state. The results show a diversity of black scores across states within regions. These are some states in nearly every region in the top, middle and bottom of the rankings. A simple regression using a SES variable for black families by state shows a positive, but insignificant result indicating that black family characteristics explains little of the distribution.[1] A state per pupil expenditure variable is positive, but insignificant.

Figure 3 shows the white distribution of scores across states. The white distribution largely follows a more familiar pattern of higher scores in the northeast and midwest with lower scores more often in the south. Exceptions occur for Texas and Rhode Island. Simple regressions here show family characteristics to be a strong predictor of the pattern of white scores, and state per pupil expenditure shows statistically significant results.

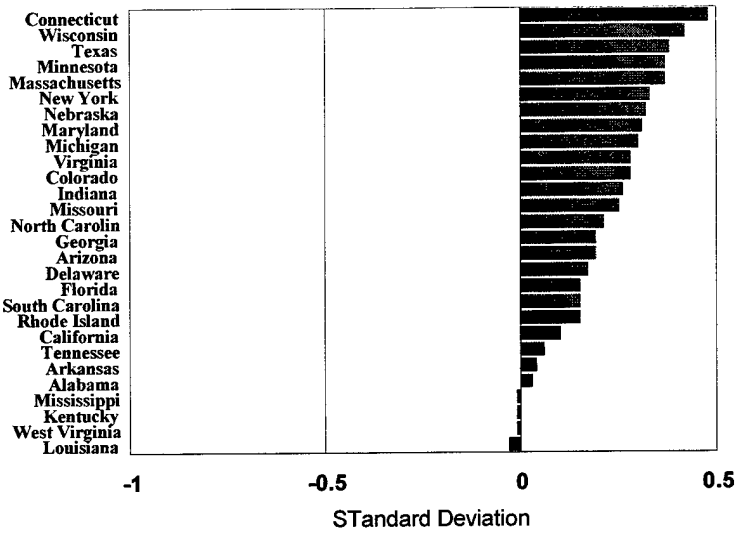
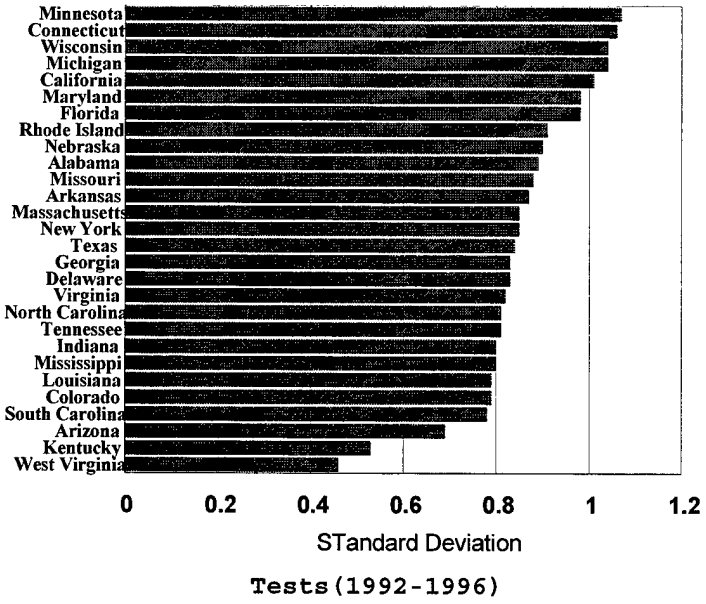


Figure 3. Average White Scores by State Across Four State NAEP Tests(1992-1996)

Figure 4 shows the black-white gap by state. The four states with the largest gap - Minnesota, Connecticut, Wisconsin and Michigan - are among the highest spending states. In each state their white students are among the highest scoring, but their black students have average to below average scores. The smallest gaps are in Kentucky and West Virginia where their white students score among the lowest state scores, but black students score above average among black student groups.

Figure 4. Average Black/White Score Gap by State Across Four NAEP



The Characteristics of Students in the Lowest Quintile

While average scores across regions and localities can tell us about differences for the average student, they do not take account of the variance within region and locality. In this section we identify the 20 percent of students (lowest quintile) scoring lowest in the nation and describe their regional, racial and locality characteristics.

Table 5 shows the racial/ethnic distribution of the lowest quintile is almost equally white, black and Hispanic. White students constitute 37.6 percent of those in the lowest quintile, while black constitute 32.8 percent and Hispanics 29.6. Although only 11 percent of white students are in the lowest quintile (last column), white students constitute 71.9 percent of all students making them the single largest group in the lowest quintile. Black and Hispanic students have much higher probabilities of being in the lowest quintile (.45 and .53 respectively), but they are much smaller proportions of the national student population (16.0 and 12.1 respectively). However, these data suggest that all races might benefit from policies directed toward low scoring students.

Race	% of total student population	% of lowest quintile population	Probability of being in the lowest quintile
White	71.9	37.6	0.11
Black	16.0	32.8	0.45
Hispanic	12.1	29.6	0.53

Table 5. Racial/ethnic characteristics of the lowest quintile

Table 6 shows similar data by region. The southeast has the largest population of low scoring students - 41 percent of low scoring students. The south and west combined have two-thirds of low scoring students. Table 7 shows the data by locality. Suburban and central city areas have about 85 percent of low scoring students with about equal numbers in each area. However, these aggregate statistics hide important variations by race within regions.

region	% of total student population	% of lowest quintile population	Probability of being in the lowest quintile
Southeast	0.33	0.41	0.25
west	0.22	0.25	0.23
Midwest	0.24	0.18	0.15
Northeast	0.21	0.16	0.16

Table 6. Characteristics of the lowest quintile by region

We now look at the location of low scoring black, Hispanic and white students. Table 8 shows where low scoring black students are located. The data on the first line for blacks in central cities in the southeast shows that they constitute 3.7 percent of the national student population and 8.5 percent of the lowest quintile population. The probability of a black student in the central cities of the south being in the lowest quintile is .46, and 23.9 percent of all students in the lowest quintile are blacks in southern central cities- the largest single group of low scoring students. Low scoring black students are highly concentrated. The last column shows that almost 80 percent of

low scoring black students live in the south or the central cities of the northeast and midwest.

locality	% of total student population	% of lowest quintile population	Probability of being in the lowest quintile
suburban	50.5	42.7	0.17
central city	29.8	42.1	0.28
rural	19.7	15.2	0.15

Table 7. Characteristics of the lowest quintile by region

Table 9 and 10 show similar data for Hispanics and white students in the lowest quintile. Hispanic low scoring students are concentrated in four areas. About 70 percent of low scoring Hispanic students are in the suburban and central cities of the south and west. Table 10 shows low scoring white students to be much less concentrated. About 54 percent are in suburbs of the south, west or midwest or the rural south.

Table 11 shows the characteristics of the lowest quintile by locality and region and provides the racial percentages of the lowest quintile within each locality/region. For instance, the suburban southeast has 14.9 percent of the total student population and 16.3 percent of the lowest quintile population. The racial/ethnic distribution of the lowest quintile students is 39.5 percent white, 43.6 percent white and 16.9 percent Hispanic. The two largest groups of low scoring students are in the suburban and central cities of the south. Approximately one in three low scoring students nationwide live in and around southern cities. Two out of three low scoring students live in the south or west. Over 80 percent are in the south, west or central cities of the northeast and midwest.

Race	locality	region	% of total student population	% of lowest quintile population	Probability of being in the lowest quintile	Distribution of lowest quintile blacks (percentage)
Black	Central city	Southeast	3.7	8.5	0.46	23.9
Black	suburban	Southeast	3.1	7.1	0.46	20.0
Black	Central city	Northeast	2.2	4.8	0.44	13.4
Black	Central city	Midwest	1.9	4.7	0.48	13.2
Black	rural	Southeast	1.2	2.7	0.47	7.6
Black	suburban	Northeast	1.4	2.4	0.34	6.8
Black	suburban	West	0.7	1.7	0.53	4.8
Black	suburban	Midwest	0.8	1.6	0.38	4.5
Black	Central city	West	0.6	1.5	0.50	4.3
Black	Rural	Northeast	0.2	0.3	0.27	0.7
Black	rural	Midwest	0.1	0.2	0.50	0.4
Black	rural	West	0.1	0.1	0.45	0.0

Table 8. Characteristics of black students in the lowest quintile

race	Locality	region	% of total student population	% of lowest quintile population	Probability of being in the lowest quintile	Distribution of lowest quintile Hispanics (percentage)
Hispanic	Suburban	West	2.5	5.1	0.42	21.9
Hispanic	Central city	West	1.8	4.5	0.50	19.3
Hispanic	Central city	Southeast	2.1	3.9	0.38	16.6
Hispanic	Suburban	Southeast	1.6	2.8	0.35	11.8
Hispanic	Central city	Northeast	1.0	2.1	0.40	8.9
Hispanic	Central city	Midwest	0.6	1.4	0.46	5.8
Hispanic	Rural	Southeast	0.6	1.2	0.42	5.3
Hispanic	Suburban	Midwest	0.6	0.8	0.27	3.3
Hispanic	Suburban	Northeast	0.7	0.7	0.20	2.9
Hispanic	Rural	West	0.4	0.5	0.28	2.3
Hispanic	Rural	Midwest	0.2	0.3	0.31	1.4
Hispanic	Rural	Northeast	0.1	0.1	0.23	0.5

Table 9. Characteristics of Hispanic students in the lowest quintile

Race	locality	Region	% of total student population	% of lowest quintile population	Probability of being in the lowest quintile	Distribution of lowest quintile whites (percentage)
White	suburban	Southeast	10.1	6.4	0.13	15.7
White	suburban	West	8.5	5.8	0.14	14.1
White	suburban	Midwest	11.0	4.9	0.09	11.9
White	rural	Southeast	5.6	4.7	0.17	11.4
White	Central city	West	4.6	3.6	0.16	8.8
White	suburban	Northeast	9.4	3.5	0.07	8.5
White	Central city	Southeast	5.1	3.2	0.13	7.9
White	Central city	Midwest	4.1	2.5	0.12	6.1
White	rural	Midwest	5.1	1.9	0.08	4.7
White	rural	West	2.9	1.8	0.12	4.4
White	Central city	Northeast	2.2	1.4	0.13	3.4
White	rural	Northeast	3.3	1.3	0.08	3.1

Table 10. Characteristics of white students in the lowest quintile

Locality	Region	% of total student population	% of lowest quintile population	% white in lowest quintile group	% black in lowest quintile group	%Hispanic in lowest quintile group
Suburban	Southeast	14.9	16.3	39.5	43.6	16.9
Central city	Southeast	10.9	15.7	20.8	54.4	24.9
Suburban	West	11.7	12.6	45.8	13.6	40.6
Central city	West	7.0	9.6	37.3	15.8	46.9
Rural	Southeast	7.4	8.6	54.0	31.5	14.5
Central city	Midwest	6.6	8.5	29.0	54.9	16.0
Central city	Northeast	5.4	8.3	16.9	57.7	25.4
Suburban	Midwest	12.4	7.2	67.3	22.0	10.7
Suburban	Northeast	11.6	6.6	52.8	36.7	10.4
Rural	West	3.3	2.5	72.7	5.8	21.5
Rural	Midwest	5.4	2.4	80.0	6.5	13.4
Rural	Northeast	3.6	1.6	77.3	15.9	6.7

Table 11. Characteristics of the lowest quintile by region, locality and race

EXPLORING ROLES FOR FEDERAL RESOURCES

A policy agenda in K-12 education must be built on a common understanding of the strengths and weaknesses of the current system, the most important new problems and opportunities that will emerge in the next 10-15 years, and a recognition that federal involvement should only occur where a clear comparative advantage exists. The federal government has traditionally played a very limited role in K-12 education providing only about seven percent of expenditures. However, there are three areas where a strong to plausible case can be made for a comparative federal advantage.

The first area - addressing the wide inequality of educational spending across states - can ***only*** be addressed by the federal government. A second area - supporting sound research and development - is a function typically centralized in public and private sector organizations. The federal government has a unique opportunity to strengthen the scientific research base in education. A more solid research base in turn could help states and districts use resources more productively. A third area where federal involvement can possibly be efficient is in increasing the supply and quality of teachers by way of incentives.

Addressing inequity in spending across states

Most of the resource inequality cannot be resolved at the state level. States spending the least are southern and western states that also have a disproportionate share of the nation's minority and disadvantaged students. Only the federal government can address this issue of interstate inequality in school spending.

Perhaps the main reason for resource shortfalls in states is their inadequate tax base. A stronger case can probably be made that the southern states lack the fiscal capacity to increase educational spending than many western states. California, Washington, Oregon and Utah all have below average expenditures, but may have more fiscal capacity to raise funds than many southern states.

Central cities pose a different problem than inequality between states. It may well be easier to address low suburban and rural scores

in the south and parts of the west than inner city scores. Central city resources depend heavily on the politics and judicial intervention within states. It may also be more difficult and more expensive to implement successful interventions in inner cities due to higher cost of living and limited space.

Besides disproportionately low spending and high numbers of disadvantaged students, there are several other reasons that urban, southern and western school districts should receive the focus of policy attention. First these students constitute a growing proportion of U.S. students, and our future productivity will depend on learning how to provide better education for them. Second, recent research suggests that the achievement scores of minority and disadvantaged students respond to additional, well-targeted educational expenditures, and that significant score gains could occur. Third, research also suggests that additional educational investment might be recouped through lower future social expenditures and improved economic productivity. Fourth, such policies would reduce the achievement gap between racial/ethnic and income groups - a source of continuing social and political divisions and economic costs in our society. Finally, improving our international standing requires lifting the scores of these students.

Addressing between state inequality by setting an adequacy level across states near the current national average spending per pupil would require a significant expansion of federal education funding - at least $25 billion annually. The largest federal K-12 education program is Title I, an $8 billion program that directs funds to low-income school districts. This program barely begins to compensate for inequalities in spending. For instance, Title I supplies about six percent of educational spending in Mississippi, but less than two percent in many wealthier states, whereas the difference in spending in Mississippi compared to wealthy states can be over 100 percent. If additional funds were effectively spent by school districts, a significant increase in scores in the south and in urban school districts could occur.

Although past evaluations of Title I have shown no compelling evidence for achievement gains, there was little guidance provided by

research and development as to how to spend the money well. For instance, teacher aides were a common use of funds, whereas current experimental research shows compelling results favoring class size reductions over hiring of aides. SO the question can future Title I like programs be more effective than past ones. Certainly, any expansion of Title I programs should limit it use to the strongest empirically supported interventions, and have built in experimental designs for evaluation. But it appears unlikely that significant progress can be made in lifting achievement nationally without significant, well structured, well targeted and well evaluated federal resources.

Federal R&D Spending on Children and Education

In most areas of national endeavor, we take progress for granted. The basis for this progress is generally good research and development that continuously produces discoveries, innovations, and new products. The nation spends about 2-3 percent of its Gross National Product on R&D in order to fuel these results. R&D in most sectors is funded through both public- and private-sector funding. The private sector funding in each area tends to provide a market-driven constraint that partially drives R&D spending to productive levels.

However, R&D in K-12 education and more broadly on children's issues is funded almost entirely by the federal government. Granted, this arrangement may be appropriate, because the public sector would be a prime beneficiary of discovery and innovation. State, local and the federal governments spend almost $600 billion a year directed toward children and youth, with education being the largest component of this spending. R&D that shows how to obtain better results with this spending could have significant payoff in savings for the public sector.

Furthermore, improving R&D on education and children could have other payoffs down the road: it could help avoid future expenditures on dysfunctional adult behaviors associated with poor educational outcomes. Much adult dysfunctional behavior appears rooted in developmental paths from childhood. Smoking, drinking, drug abuse, criminal behavior, depression and other mental health conditions, job performance and stability and even factors linked to heart disease have roots in

behavior and lack of educational success established while growing up. R&D that begins to understand and address these issues may decrease the substantial societal costs of these behaviors for adults. Investment in the first 20 years of life may have substantial impact on the quality of the last 20 years of life, and easily recoup any additional investment required.

Currently, only about 0.3 to 0.4 percent of public spending on children and education is devoted to R&D--a far smaller proportion than the average 2-3 percent spent in other areas. This R&D represents only about 5 percent of total domestic federal R&D, and about 2 percent of national R&D (Office of Science and Technology, 1997). Since almost all R&D on children's issues is federally funded, it is problematical whether R&D seeks the most productive level. Evidence from four other domestic areas would suggest systematic underinvestment by the public sector in R&D for functions that, like K-12 education, are almost entirely public-sector responsibility: air traffic control systems, tax collection systems, postal systems and voting systems. Each of these systems is in crisis partly because of failure of R&D to produce innovations and improved products and the slow adoption of new technology. If the public sector is inefficient in allocating R&D in areas where it bears the prime responsibility, perhaps it is not surprising that progress in education and improving outcomes for our children has been slow and uncertain.

Past research in education has been notable for its inconsistency and inability to provide compelling results. Under such circumstances, low funding might have been appropriate. There is emerging evidence that future R&D on education and children could be more scientifically productive. Two central concepts underlie these assertions.

First, a restructuring of educational research around the model used in health research would improve its scientific quality. Health research--widely hailed for advancing knowledge and producing new discoveries, innovations, and technology-- is undergirded by solid basic and applied research based on the use of clinical trials and of major ongoing longitudinal studies tracking health behavior and status focusing on virtually every health problem.

Health research also has an institutional infrastructure based on teaching hospitals, schools of public health and academic research centers that closely link research to training, practice and public education. Another component of this infrastructure is extensive cross-training among disciplines and between researchers and practitioners. All of this is aided by a central federal funding agency--the National Institutes of Health--that is guided by scientific peer review and is able to set priorities and achieve more efficient investment in infrastructure.

By contrast, research on children outside physical health often lacks a vibrant basic research component, has almost no scientifically structured clinical trials, has relatively few major longitudinal surveys tracking children and no equivalents of teaching hospitals or schools of public health that combine research with practice. The R&D community is fragmented across disciplines and federal departments that invest in infrastructure, making investment inefficient. Research centers tend to be university-based and far removed from schools or communities where children's learning and development occurs. Research on children can strengthen its scientific basis by moving toward the health model.

Second, new opportunities for breakthroughs in understanding children's cognitive development are emerging in current brain research. This work promises to illuminate the basic processes underlying cognitive, emotional and physical development in children. Just as physics and chemistry needed understanding of atomic and nuclei structure and electron movement in solids to spawn revolutions in R&D and technology, and medicine and biology needed access to cells and genes to move forward, research on children needs a basic understanding of the organ that centrally governs all development in children- the brain. Each of these past revolutions occurred with advances in instrumentation that allowed deeper or more precise penetration into matter, i.e., particle accelerators, electron microscopes, lasers and radio telescopes.

The new brain instrumentation promises to provide the same penetration allowing increasingly precise knowledge of the links between

brain processes and behavior. This knowledge will ultimately spawn a revolution in understanding and the related capacity to address developmental problems in children.

A strong scientific infrastructure must emerge that links this basic brain research to more applied areas of cognitive, emotional, physical and social development and to training and practice. It takes at least 10 years to substantially restructure areas of research. Such restructuring has occurred in the area of health in the 1960s and 1970s, in computer science when it emerged from the field of electrical engineering in the same period and engineering in the last 15 years. Federal funding was instrumental in nurturing and building these new fields through support for research, funding for new undergraduate and graduate programs and writing of new textbooks and curriculums.

The federal government needs to develop a longer-term strategy for accomplishing this restructuring to shape and guide research on K-12 education and children. Some in Congress have recommended doubling budgets in NIH and NSF over 5 years, which would increase R&D spending alone by over 20 billion dollars. However, only a small portion of this—about a billion dollars--would naturally flow to children's R&D. However, a small increase in the portion of scheduled federal R&D increase toward children's research could double R&D spending on children-currently at around 3 billion dollars. Such increases should occur gradually and only with a plan that insures better research productivity.

Improving Teacher Quality

Perhaps the most important ***emerging*** issue in education amenable to federal involvement is the potential shortage of certified and qualified K-12 teachers in the next 10 years. The demand for new teachers is predictably increasing due to a disproportionately large proportion of teachers nearing retirement, increasing demand from smaller class size, rising enrollments and rising teacher attrition levels due partly to a highly competitive labor market (Hussar, 1999). Expanding the supply of new teachers to meet rising demand may be difficult due to the highly competitive labor market for new college graduates and non-competitive

wage for teachers. Shortages may be particularly severe for science and math teachers where the gap between wages inside and outside teaching are very large.

Shortages of certified teachers are usually addressed through increasing teacher salaries at the school district level and/or hiring of uncertified teachers. Unfortunately teacher salaries are raised by increasing salaries for all teachers- an inefficient and very expensive process since it is salaries for beginning teachers that is most closely linked to decisions by college students to teach. Since teacher salary and benefits consume over 50 percent of educational expenditures, rising salaries for all teachers could consume a considerable portion of the growth in education expenditures over the next decade leaving little for implementing more effective reform initiatives.

When teacher shortages occur, they fall disproportionately on inner city schools and lower paying school districts- predominately in the south. These areas not only will need to hire more teachers to meet their rising demand, but also will need to replace many current teachers who will take jobs that open up in higher paying districts. Inner city teachers will take suburban jobs, and inter state competition may drive teachers north. Thus, southern and urban school districts will be hardest hit- exactly those districts that constitute the major problem areas for U.S. education.

The federal government can efficiently increase teacher supply and raise the quality of incoming teachers through several programs aimed at beginning and younger teachers. These include college loan forgiveness programs, undergraduate scholarships and a graduate level "GI Bill" type program. All of these type of programs have been used to increase the quality of officer and enlisted personnel in our Armed Forces. A contributory "GI Bill" would allow the federal government to provide matching funds that could be used for pursuit of advanced degrees after a specified number of years of teaching. This program would mainly appeal to college grads with plans for graduate school, but uncertain of career choice or lacking funds. Such candidates are likely to be of higher quality on average than typical college students choosing teaching.

Such programs can also be more heavily targeted toward types of teachers in short supply (math and science, inner city) or teachers with superior grades or credentials(a math degree rather than education degree), thereby increasing even more the quality of incoming teachers. All of these programs essentially provide the equivalent of more front-end compensation for younger teachers that does not get built into the entire salary scale- making it efficient. These programs could ease the need for across the board salary increases and leave more discretionary funds at the state and local level for effective initiatives- especially in urban and southern schools. The long term cost of such programs that would on average provide the equivalent of a 10 percent salary increase for younger teachers would be around 2-3 billion dollars annually.

BIBLIOGRAPHY

Campbell, J.R., Donahue, P.L., Reese, C.M., and Phillips, G.W. NAEP 1994 Reading Report Card for the Nations and the States; Findings from the National Assessment of Educational Progress and Trail State Assessment, Washington, DC: National Center for Education Statistics, 1996.

Evans, William N., S. Murray and R. Schwab, The Impact of Court-Mandated School Finance Reform, in Equity and Adequacy in Education Finance, National Research Council, Helen Ladd, Rosemary Chalk and Janet Hansen, National Academy Press, 1999

Evans, William N., Sheila Murray, and Robert Schwab, 1997. "Schoolhouses, Courthouses, and Statehouses after Serrano." Journal of Policy Analysis and Management 16 (January): 10-31.

Ferguson, Ronald F. and Helen F. Ladd, "How and Why Money Matters: An Analysis of Alabama Schools." In Helen F. Ladd, ed., 1996. Holding Schools Accountable Washington, D.C.: The Brookings Institution, 1996.

Ferguson, Ronald F., "Can Schools Narrow the Black-White Score Gap," in The Black-White Score Gap, in The Black-White Score Gap, C. Jencks and M. Phillips, eds., Brookings, Wash D.C., 1998a

Ferguson, Ronald F., "Teachers' Perceptions and Expectations and the Black-White Score Gap" in The Black-White Score Gap, C. Jencks and M. Phillips, eds., Brookings, Wash D.C., 1998b

Finn Jeremy and C. Achilles, Tennessee's Class Size Study: Findings, Implications and Misconceptions, Educational Evaluation and Policy Analysis, Summer, 1999, Vol 20, No.2

Greenwald, Rob, L.V. Hedges and R. Laine, The Effect of School Resources on Student Achievement, Review of Educational Research, Fall, 1996 Vol 66, No 3, pp. 361-396

Grissmer, David, A. Flanagan and S. Williamson, "Does Money Matter for Minority and Disadvantaged Students: Assessing the New Empirical Evidence, in Developments in School Finance: 1997, William Fowler(ed.), U.S. Department of Education, NCES 98-212, 1998c

Grissmer, David,W., Ann Flanagan, Jennifer Kawata and Stephanie Williamson, Improving Student Achievement: What Do State NAEP Scores Tell Us, RAND, 2000

Grissmer, David, Assessing the Evidence on Class Size: Policy Implications and Future Research Agenda, Educational Evaluation and Policy Analysis, Summer, 1999, Vol 20, No.2

Grissmer, David W. and Ann Flanagan, Moving Educational Research Toward Scientific Consensus, in Analytic Issues in the Assessment of Student Achievement, David Grissmer and Michael Ross (eds.), forthcoming

Hanushek, Eric A., "The impact of differential expenditures on school performance," Educational Researcher, Vol. 18, No. 4, pp. 45-51, 1989.

Hanushek, E. A., Making Schools Work: Improving Performance and Controlling Costs, Washington, DC: The Brookings Institution, 1994.

Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor, "Aggregation and the Estimated Effects of School Resources," The Review of Economics and Statistics, (Fall), pp. 611-627, 1996.

Hanushek, Eric, Assessing the Empirical Evidence on Class Size Reductions from Tennessee and Non-Experimental Research, Educational Evaluation and Policy Analysis, Summer, 1999, Vol 20, No.2

Hedges, Larry V., Richard D. Laine, and Rob Greenwald, "Does Monday matter: Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes," Educational Researcher, Vol. 23, No. 3, 1994, pp. 5-14.

Hedges, Larry V., and Amy Nowell, "Group Differences in Mental Test Scores: Mean Differences Variability, and Talent," in C.S. Jencks and M. Phillips (eds.) The Black-White Test Score Gap, forthcoming 1998.

Jencks, Christopher and Meredith Phillips, The Black-White Test Score Gap, Brookings, Wash D.C., 1998

Krueger, Alan B., "Reassessing the View That American Schools Are Broken," Economic Policy Review, March, 1998, pp. 29-43.

Krueger, A. B. (1999). Experimental estimates of education production functions. Quarterly Journal of Economics, CXIV, 497-532.

Ladd, Helen F., and Janet S. Hansen, Making Money Matter: Financing America's Schools, National Academy Press, 1999, Wash D.C.

Ladd, Helen, Rosemary Chalk and Janet Hansen, Equity and Adequacy in Education Finance, National Research Council, National Academy Press, 1999

Molnar, Alex et al, Estimated Achievement and Teacher Time Allocation Effects from a Quasi-Experimental Class Size Reduction in Wisconsin, Educational Evaluation and Policy Analysis, Summer, 1999, Vol 20, No.2

Nye, B. A., Hedges, L., & Konstantopoulos, S. (1999a). The effects of small class size on academic achievement: The results of the Tennessee class size experiment. Manuscript under review, Department of Education, University of Chicago.

Nye, Barbara, Larry V. Hedges and Spyros Konstantopoulos, The Long Term Effects of Small Classes: A Five-Year Follow-up of the Tennessee Class Size Experiment, Educational Evaluation and Policy Analysis, Summer, 1999b, Vol 20, No.2

Parrish, Thomas B., and Christine Hiikido, Inequalities in Public School District Revenues, NCES 98-210, National Center for Educational Statistics, 1998

This paper was prepared for the Brookings Institution Conference on the Black-White Test Score Gap, Washington, D.C., February 1-2, 2001. The work reported herein was supported under the Educational Research and Development Centers Program PR/Award No. #R902F98001 administered by the Office of Educational Research and Improvement, U.S. Department of Education. However, no official support by the Department of Education is intended or should be inferred.

[1] We used a SES variable constructed from 1990 Census characteristics for black families by state weighted by coefficients from equations derived from the National Educational Longitudinal Study(NELS). The variable includes parental education, family income, family size, average age of mother at child's birth and single parent. Grissmer et al, 2000 has a more detailed derivation in Appendix E.